(12) UK Patent Application (19) GB (11) 2 202 706 (13) A

(43) Application published 28 Sep 1988

(21) Application No 8707423

(22) Date of filing 27 Mar 1987

(71) Applicant
British Broadcasting Corporation

(Incorporated in United Kingdom)

Broadcasting House, London, W1A 1AA

(72) Inventor
Timothy John Borer

(74) Agent and/or Address for Service
Reddie & Grose
16 Theobalds Road, London, WC1X 8PL

(51) INT CL⁴
H04N 5/213 H04N 3/36

(52) Domestic classification (Edition J):
H4F D12K9 D12X D1B9 D22 D30K D30T3 D30X D3 D54 HD

(56) Documents cited
None

(58) Field of search
H4F
Selected US specifications from IPC sub-class H04N

(54) Video signal processing

(57) An input video signal is delayed by two cascaded picture delays 10, 11. The undelayed and 2-picture delayed signals are processed by interpolators 13, 14 to produce compensated signals in which moving objects are subjected to such relative delays that the moving objects register in the two compensated signals and the 1-picture delayed signal. These three signals are fed to a median selector 12 which selects the median value and thereby temporally filters the video signal in such a way as to remove impulse noise, e.g. arising from scanning dirty film. The interpolators 13, 14 are controlled in dependence upon motion vectors describing movement for each pixel and derived from the input signal by a motion measurement unit 16 or received along with the input video signal in a digitally assisted television receiver.

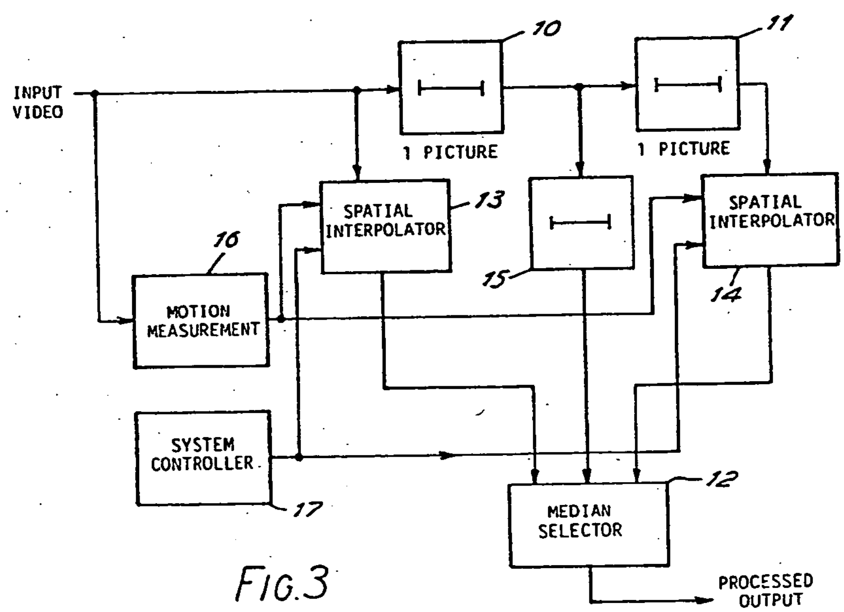

FIG.3

GB 2 202 706 A

The drawing(s) originally filed was (were) informal and the print here reproduced is taken from a later filed formal copy.

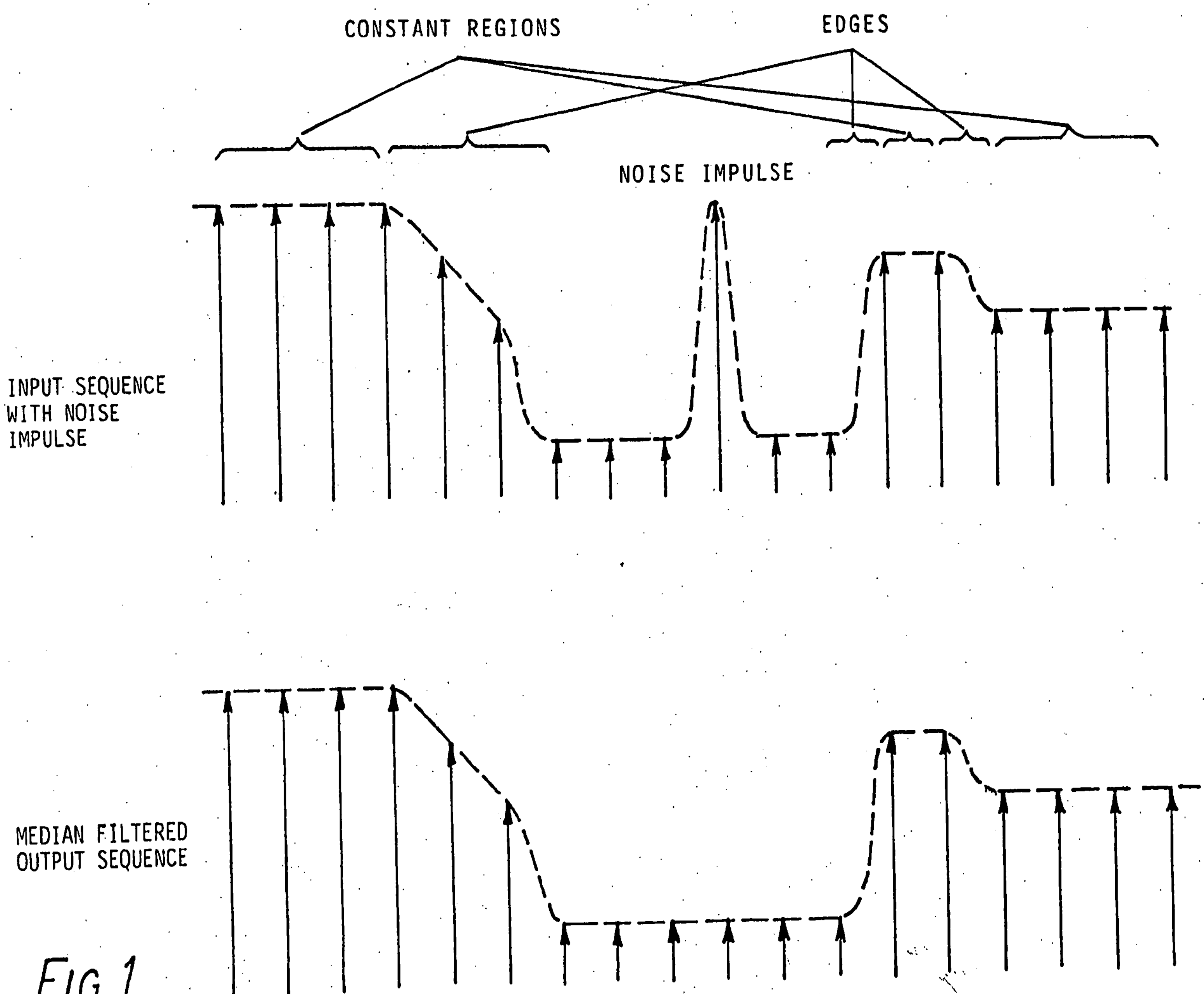
CONSTANT REGIONS
EDGES
NOISE IMPULSE
INPUT SEQUENCE
WITH NOISE
IMPULSE
MEDIAN FILTERED
OUTPUT SEQUENCE

FIG. 1

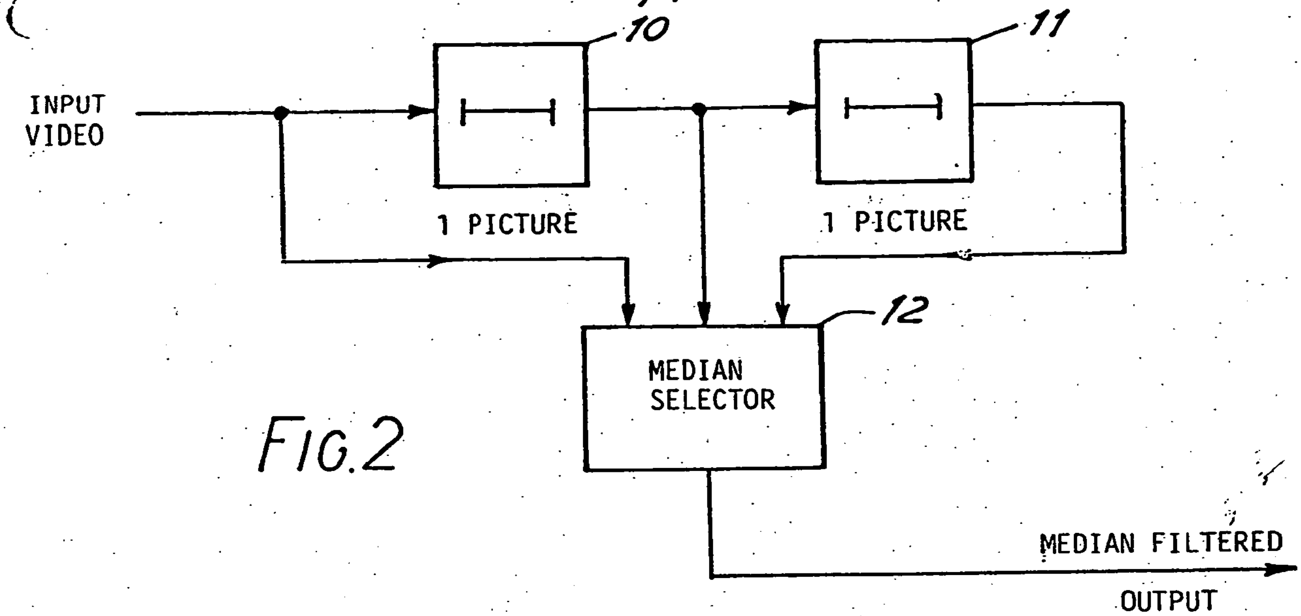
10
11
INPUT VIDEO
1 PICTURE
1 PICTURE
12
MEDIAN SELECTOR
MEDIAN FILTERED OUTPUT

FIG. 2

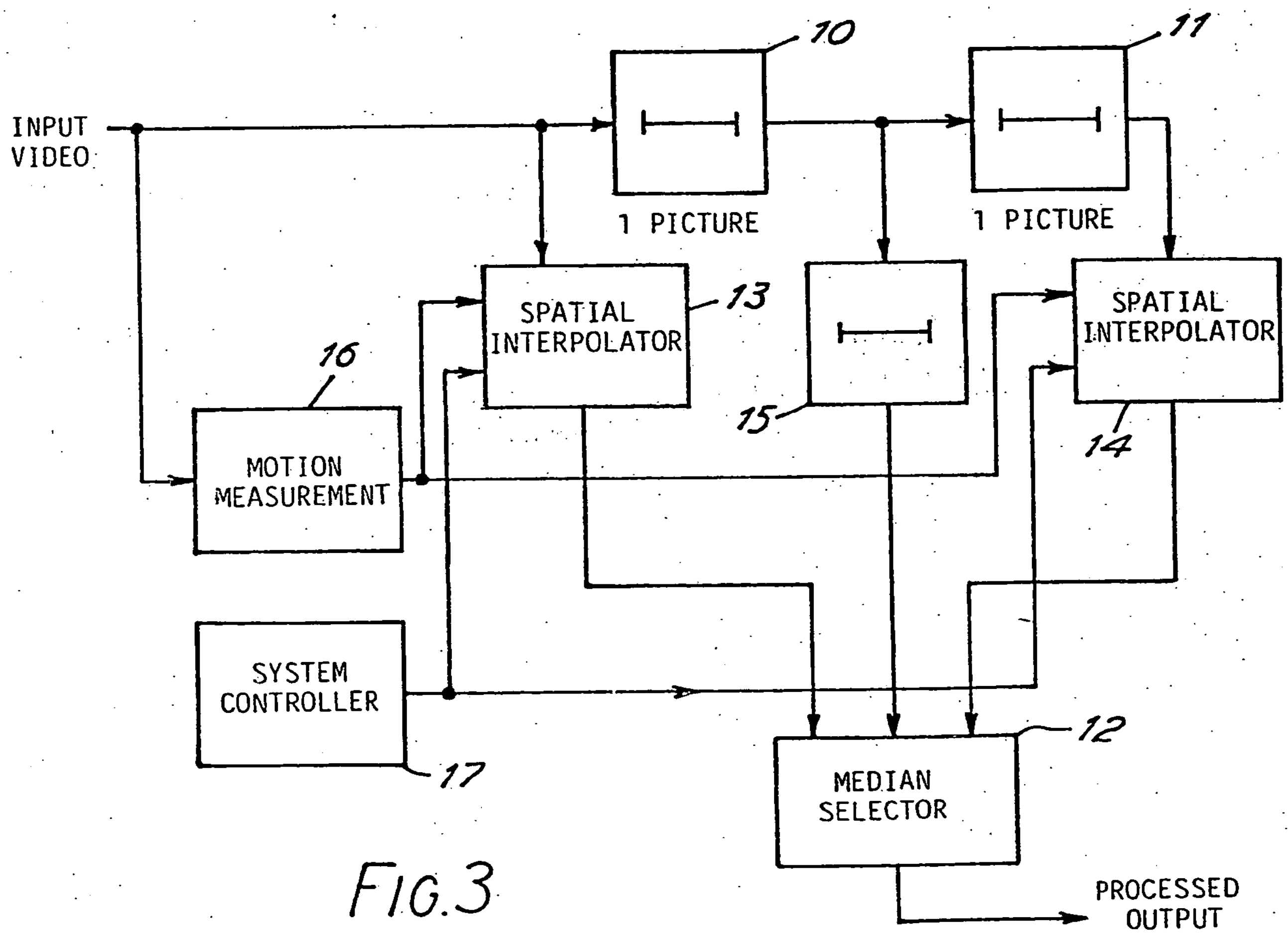
10
11
INPUT VIDEO
1 PICTURE
1 PICTURE
SPATIAL INTERPOLATOR
13
SPATIAL INTERPOLATOR
14
15
16
MOTION MEASUREMENT
SYSTEM CONTROLLER
17
12
MEDIAN SELECTOR
PROCESSED OUTPUT

FIG. 3

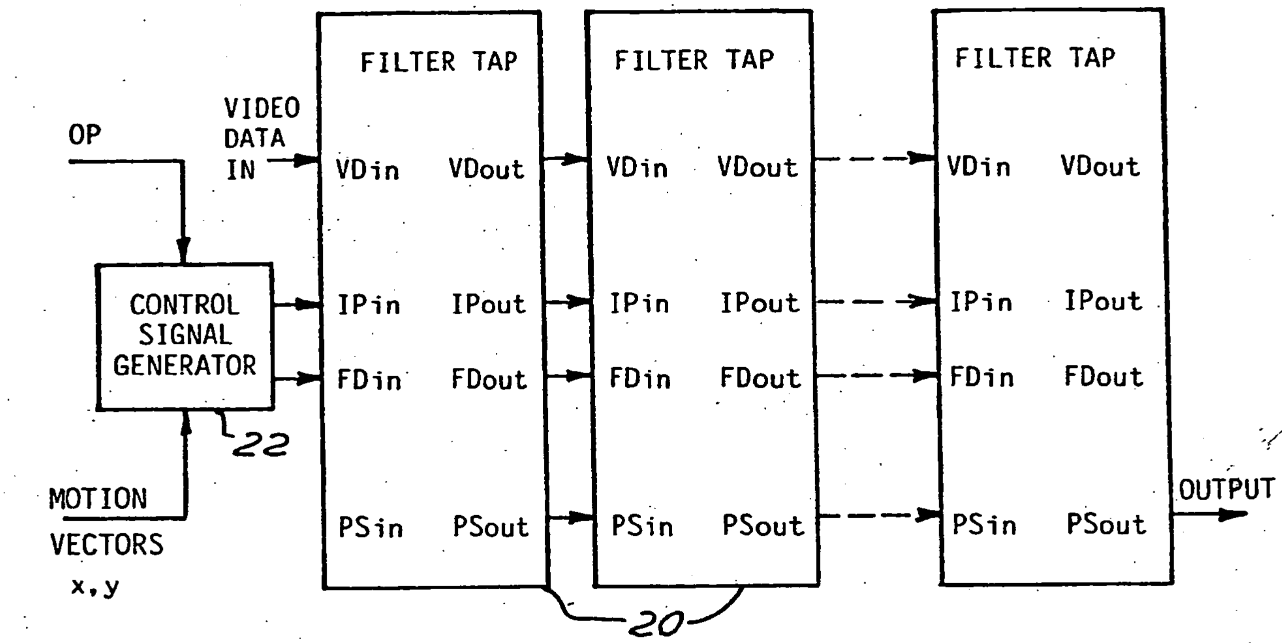
OP
VIDEO DATA IN
CONTROL SIGNAL GENERATOR
22
MOTION VECTORS x, y
FILTER TAP
VDin
VDout
IPin
IPout
FDin
FDout
PSin
PSout
FILTER TAP
VDin
VDout
IPin
IPout
FDin
FDout
PSin
PSout
FILTER TAP
VDin
VDout
IPin
IPout
FDin
FDout
PSin
PSout
OUTPUT
20

FIG.4

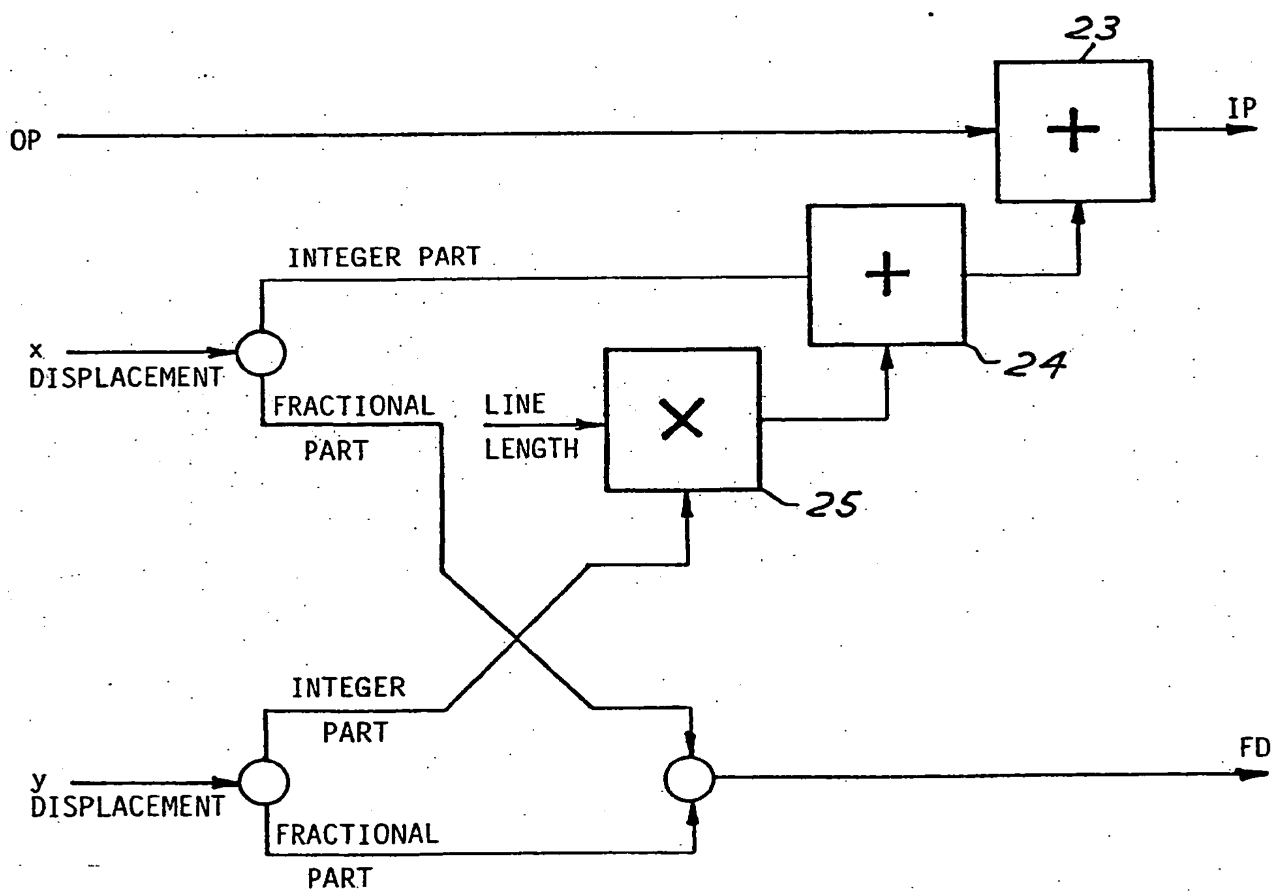
OP
23
+
IP
INTEGER PART
+
24
x DISPLACEMENT
FRACTIONAL PART
LINE LENGTH
×
25
INTEGER PART
y DISPLACEMENT
FRACTIONAL PART
FD

FIG.5

2202706

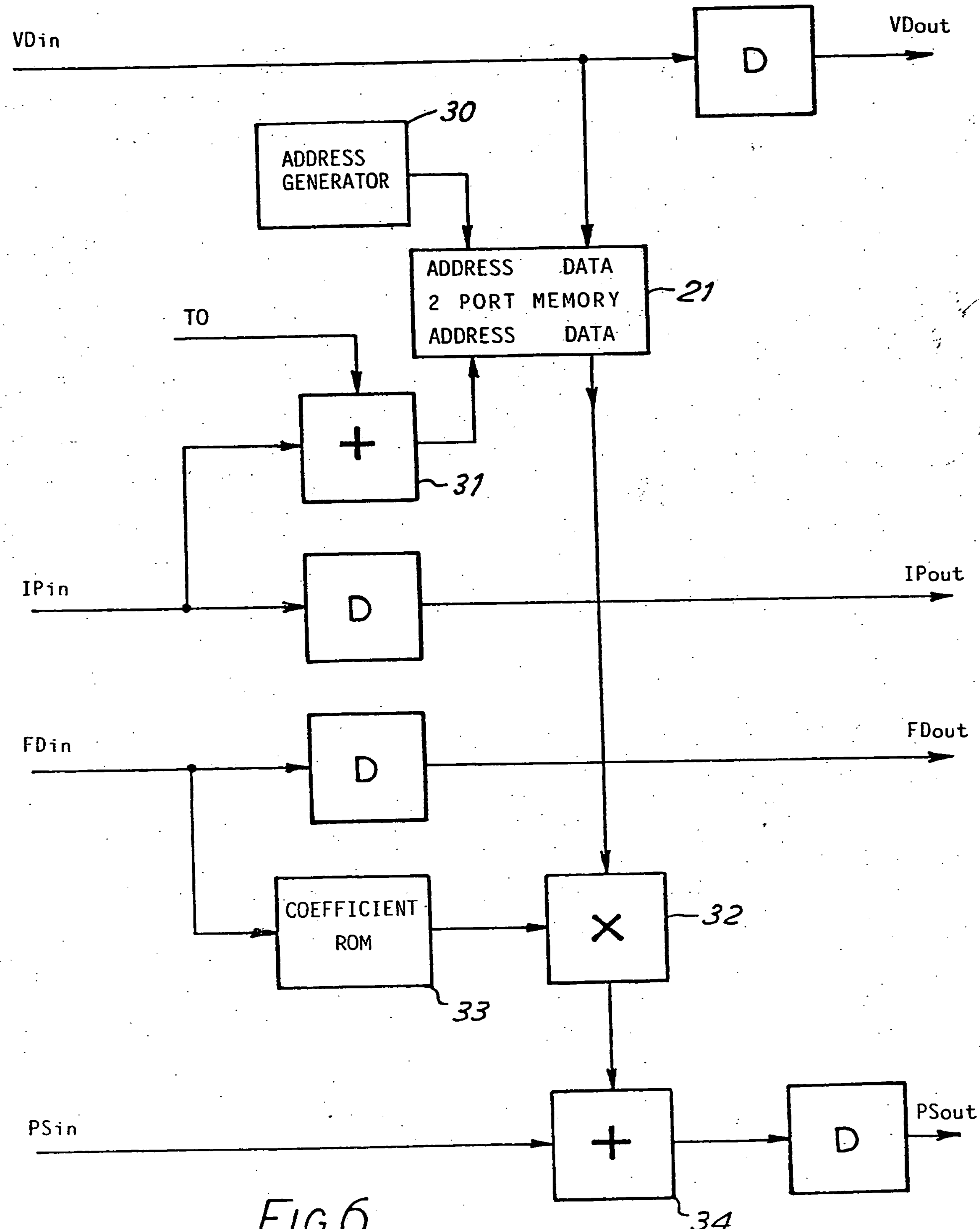

FIG.6

VIDEO SIGNAL PROCESSING

The present invention relates to the processing of video signals to remove impulse noise, particularly noise arisng when a dirty film is scanned, and when the signals represent moving images.

We have already proposed a method of measuring the motion in television pictures in our International Patent Application PCT GB86/00796. It has been shown that it is possible to measure motion precisely and reliably. Availability of such motion information enables the use of new signal processing methods.

The present invention relates to a method which will be referred to as motion compensated median filtering. Applications of this method include improved concealment of film dirt and DBS interference (using DATV). In addition about 3dB of noise reduction can also be achieved. The invention also relates to apparatus for use in performing the method.

DATV stands for Digitally Assisted Television and relates to a technique described in our International Patent Application PCT GB86/00799 in which a digital signal accompanying the television signal contains supplementary information, such as motion vector information.

Programs from film form a substantial proportion of broadcasters output. This is likely to remain true for the foreseeable future, if only because of feature film and archive material. One of the major impairments of film is the presence of film dirt. Two methods of dirt concealment are currently available. One requires a specially designed telecine, the other is purely electronic. Both methods have advantages and drawbacks; neither is perfect. In both techniques the dirt is first detected, then an algorithm is used to conceal it. By contrast median filtering combines the two processes.

The first technique (Childs, I. February 1985. Further developments of CCD line array telecine. BBC Research Department Report No. BBC RD 1985/3) detects dirt by its infra red absorption. Fortunately colour film emulsions are transparent to infra red. Thus infra red absorbtion can be assumed to be due to dirt. The

technique cannot detect dirt on black and white film because the silver image is opaque to infra red. Neither can it detect the printed image of dirt on the negative. Limitations of lens performance at infra red make detection less effective for small rather than large dirt. Infra red dirt detection could be integral to new solid state telecines. Practical problems, however, prevent its incorporation into existing machines.

The electronic method of dirt detection (Storey, R. February 1985. Electronic detection and concealment of film dirt. BBC Research Department Report RD 1985/4) has largely complementary characteristics. It can readily detect small positive or negative dirt on monochrome or colour film. Unfortunately occasional motion impairment results if it is used to detect large dirt. It is also limited to detecting small dirt in moving areas. Subjectively, however, large dirt is very significant.

With both techniques simple algorithms are used to interpolate the corrupted data. With electronic dirt detection, dirt is replaced by the average of the two adjacent pictures. This technique was found not to work with infra red dirt detection, the reason being that a temporal average is an inappropriate substitution in a moving area. With large dirt in moving areas this inappropriate substitution can be almost as objectionable as the dirt. For infra red dirt detection a simple spatial interpolation was used.

Impulse noise in a DBS receiver is similar in nature to film dirt. So techniques which conceal film dirt should also conceal such impulse noise. If the receiver had to incorporate a complete motion compensated median filter it would be uneconomic. If however the receiver already incorporates DATV bandwidth reduction the additional cost would be negligible.

The object of the present invention is to provide a method and apparatus which will filter a signal more effectively than the known methods described above, combining the advantages but eliminating the weaknesses of the known methods.

The method and apparatus according to the invention are defined in the claims.

The invention will be described in more detail, by way of example, with reference to the accompanying drawings, in which:

Fig. 1 is a diagram used in explaining 3-point median filtering,

Fig. 2 is a block diagram of a temporal median filter,

Fig. 3 is a block diagram of a motion compensated temporal median filter,

Fig. 4 is a block diagram of one spatial interpolator of Fig.3,

Fig. 5 is a block diagram of a control signal generation for the spatial interpolator, and

Fig. 6 is a block diagram of one tap of the spatial interpolator.

Television pictures are moving images sampled in 2 or 3 dimensions. For the purposes of digital signal processing they are sampled in 3 dimensions (horizontal, vertical and temporal). Unless the three sampling rates are sufficiently high, information about the moving image will be lost. Spatially, sampling of television pictures is (almost) sufficient. Temporally, however, television and film images are greatly under sampled. This temporal undersampling results in aliasing of the signal spectrum.

Viable signal processing algorithms must allow for temporal undersampling. Failure to do this will result in unacceptable impairments to the processed picture. To make allowance for temporal undersampling additional a priori information is required about television pictures. A frequently made assumption is that much temporal aliasing is due to movement. A motion detector is then used to switch between two algorithms, one for stationary and one for moving areas.

Motion compensation assumes that the image consists of a number of independently moving objects. This model is expressed mathematically in equation 1.

$$g'(x,y,t)=SUM[g_i(x-u_i t,y-v_i t)] \qquad \text{Equation 1}$$

g' is the composite image and g_i are the component objects moving with velocity $(u_i v_i)$. This model, while not exact, is a good approximation to reality. Each pixel within an image is assigned

its own motion vector., This indicates in which direction the pixel is moving. The motion vector is used to provide the correct motion compensation required by each pixel.

The most obvious limitation of equation 1 is that it makes no allowance for covered or uncovered background. This should not present a serious problem for several reasons. Firstly the area of uncovered background is quite small. Secondly the eye takes a finite time to resolve a new image; by this time uncovered background has become part of an object. Finally any problems which do arise can be solved by determining what the uncovered background should be from several adjacent pictures.

The model of equation 1 assumes linear, rigid body motion. Temporal frequencies in a moving image can arise from sources other than linear motion. These sources include, non-linear motion (acceleration), changes of shape of objects and rotations in any of three spatial dimensions. These factors can be allowed for in actual processing by permitting some temporal frequencies other than those resulting from pure linear translation.

Motion compensated processing treats the component objects within an image as if they are stationary. Each pixel (or block) in an image is processed in a frame of reference moving at the same 'velocity' as that pixel. This is achieved by shifting objects in adjacent picture frames so that they coincide spatially. This, in turn, requires each image to be spatially interpolated to allow for non-integer pixel shifts. This is described in our International Patent Application PCT GB86/00795 and in Thomas, G.A. October 1986. Bandwidth Reduction by Adaptive Subsampling and Motion Compensation DATV Techniques. 128th SMPTE Technical Conference, October 24-29 1986, New York. Storey, R. June 1986. HDTV Motion Adaptive Bandwidth Reduction using DATV. BBC Research Department Report No. BBC RD 1986/5.

The need for spatial interpolation may require motion compensated processing to perform implicit interlace to sequential conversion. Transparent interlace to sequential conversion may be impossible. Acceptable results can, however, be achieved (for example, by again utilising motion information). In this description, a sequential source has been assumed. Where this is

not applicable, i.e. for video rather than film sources, it is assumed that the necessary interlace to sequential conversion has been performed.

Median filtering is a nonlinear signal processing technique useful for removing impulse noise while preserving edges. Such filters are easy to implement digitally and effective in practice. Unfortunately they are difficult to analyse mathematically. Nevertheless a considerable body of theory has been developed, see the following references:

Wendt, P.D., Coyle, E.J., Gallagher, N.C. March 1986. Some convergence properties of median filters. IEEE Transactions on Circuits and Systems, Vol. CAS-33, No.3 p.276-286.

Fitch, J.P., Coyle, E.J., Gallagher, N.C. February 1985. Root properties and convergence rates of median filters. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-33, No.1 p.230-239.

Fitch, J.P., Coyle, E.J., Gallagher, N.C. December 1984. Median filtering by threshold Decomposition. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-32, No.6 p.1183-1188.

Gallagher, N.C., Nodes, T.A. October 1982. Median Filters : Some modifications and their properties. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-30, No.5 p.739-746.

Gallagher, N.C., Wise, G.L., December 1981. A theoretical analysis of the properties of median filters. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, No.6 p.1136-1141.

One dimensional median filtering consists of replacing the centre pixel in a window by the median of the pixels within that window. Thus median filtering is the calculation of a running median for a sequence of pixels. Linear filtering, by comparison, is the calculation of a weighted running mean for a sequence of pixels. For ease of implementation the window usually encompasses an odd number of pixels. However this is not essential. See Kendal, M., Stuart, A. The advanced theory of statistics: Volume 1 Distribution theory. Fourth edition 1977. Charles Griffin &

Company Ltd. ISBN 0 85264 242 3, p 39-40, for a definition of the median.

An example of median filtering is shown in Figure 1. The top sequence is filtered using a three point median filter. The output is the bottom sequence. Note that the impulse noise is completely removed while edges and constant regions are preserved. Linear filtering, in the same example, would not only have failed to completely remove the noise, it would also have smoothed the edges.

The condition for removal of noise is that the noise should be impulse noise and the wanted signal should be a root. A root signal is invariant to median filtering. It can be shown (the first article by Fitch et al mentioned above) that a sufficient condition for a signal to be a root is that it consist of constant regions and edges only. For a filter of length N a constant region is defined as at least N-1 consecutive, identically valued points. An edge is defined as a monotonic region between constant regions of different value.

Other nonlinear filtering techniques are available. These include, for example, the modified trimmed mean filter as described in Lee, Y.H., Kassam, S.A. June 1985. Generalised median filtering and related nonlinear filtering techniques. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-33, No.3 p.672-683. However the median filter has been described as it is simple and effective.

It is proposed to apply median filtering in the time dimension to remove dirt and other artifacts. For non motion compensated filtering this would be achieved by regarding the moving image as a set of one dimensional signals. Each one dimensional signal consists of the sequence of points at a particular spatial location and would be median filtered to produce an output for that location. A hardware implementation of this is shown in Figure 2.

The input video signal is passed through two picture delays 10 and 11 and the three video signals thereby made available are the inputs to a median selector 12 which processes the three inputs in a manner analogous to that explained with reference to Fig 1.

Motion compensated filtering is achieved by varying the delays 10 and 11 in Figure 2 on a pixel by pixel basis. This effectively

aligns corresponding points on a moving object. Moving objects can then be treated as if they were stationary. Objects are thus processed in their own, moving, frame of reference rather than in the stationary frame of the complete image.

For proper motion compensation the delays in Figure 2 must be variable to less than one pixel. In practice this means that the delay elements must also include a spatial interpolator. The need for precisely variable delays is because the motion of objects is continuously variable. Therefore in order to align a moving object in two pictures, continuously variable delays are required.

An implementation of motion compensated median filtering is shown in Figure 3. In addition to the picture delays 10 and 11 and median selector of Fig. 2, the input video is processed by a spatial interpolator 13 and the video signal delayed by two pictures is processed by a spatial interpolator 14. The video signal delayed by one picture passes through an additional fixed delay 15 which compensates for the delays in the interpolators.

It is assumed that the motion vectors required to control the interpolators 13 and 14 are derived locally by a motion measurement circuit 16, as described in the first International Application referred to above. However, the motion vector information could be received from the transmitter in a DATV receiver. A system controller 17 sequences the operations of the interpolators. The median selector is very simple, consisting of a few comparators and registers whereby the three inputs are compared and their median is selected as the processed output.

The most hardware intensive parts of a motion compensated median filter are the spatial interpolators. These are required to implement the variable sub-pixel shifts. A block diagram of a spatial interpolator is given in Figure 4. It consists of multiple filter taps cascaded in a ladder structure. The constituent filter taps are illustrated in Figure 6. Additional delays required for pipelining are not shown.

The spatial interpolator (Fig 4) consists of cascaded filter taps 20, each of which takes an input partial sum PS and adds the required increment thereto, the last tap providing the

interpolation output. PSin for the first tap is zero. The required increments are derived from video data VD which is passed along the taps. Each tap (Fig.6) includes a 2-port memory 21 storing at least partof a picture. Ideally a complete picture is stored to enable the system to cope with any range of vertical movement but a smaller store will suffice. For example a 50-line store would provide for vertical motions of +/-20 lines per frame with a vertical filter aperture of 10 taps.

The system controller 17 (Fig.3) furnishes the addresses of required output pixels corresponding to the motion vectors provided by the unit 16. Each such vector consists of an x (horizontal) displacement and a y (vertical) displacement. A control signal generator 22 (Fig.4) computes the required input pixel address IP and a fractional displacement FD derived from the fractional part of x and y. IP and FD are also passed along the taps 20.

The address of the input pixel corresponding to a given output pixel OP with a motion vector x, y is readily seen to be the output pixel address plus INT(x) plus INT(y) LINE LENGTH where INT( ) represents integral part of, the pixel address space is linear and LINE LENGTH is the number of pixels in one television line. This algorithm is implemented by adders 23, 24 and a multiplier 25 in Fig.5.

The fractional parts of x and y form a two component vector FD.

Referring now to Fig. 6, various delays D are compensating delays maintaining the required synchronism between the outputs VD, IP, FD and PS. The incoming video data is written into the picture memory 21 under control of a cyclic address counter 30 synchronized to the VD in signal at the tap. Data is read out of the memory by the sum (adder 31) of IPin and a filter tap offset TO which will be explained below. The data read out is multiplied in a multiplier 32 by a coefficient from a coefficient ROM 33 addressed by FD and the resulting increment is added to PSin by an adder 34 to form PSout.

The flexibility required by a motion compensated interpolator is provided by the 2-port memory in Figure 6. This effectively generates a variable delay by changing the memory read out address. Moving objects in different pictures can thus be aligned prior to temporal filtering.

Output pixels from the spatial interpolator can be calculated in any order. The pixel which is calculated is determined by the input signal OP specifying the output pixel address (see Figure 4). Each filter tap adds a contribution corresponding to a different region of the two dimensional impulse response.

Each spatial interpolator tap 20 is customised to correspond to a particular region of the complete impulse response. This is achieved by programming the coefficient ROM and setting the filter tap offset TO (see Fig.6). The ROM 33 is programmed with the impulse response within the chosen region. The filter tap offset TO determines the effective delay generated by the 2- port memory. Contributions from different regions of the impulse response are accumulated as the signals flow through the interpolator.

Motion compensated median filtering has been conceived as a method of concealing film dirt. To be of any use it must demonstrate considerable advantages over the techniques which are currently available and described above.

Median filtering will detect both positive and negative dirt of any size on monochrome or colour film. It thus combines the best features of the two known methods. The dirt is replaced by appropriately shifted (allowing for movement) data from an adjacent picture. The size of the dirt does not matter as the algorithm processes each pixel separately. The maximum subjective improvement can therefore be achieved by concealing both large and small dirt.

Motion compensated median filtering should introduce a minimum of artefacts. The assumption for dirt concealment is that the image of an object is a root signal. The image of an object only undergoes small changes from picture to picture so this is likely to be true. No special precautions are required for shot changes. Since these constitute an edge in the temporal signal they are passed, unchanged, by a median filter.

While this technique was originally conceived for dirt concealment it may have other applications. For example impulse noise in DBS Transmissions has similar characteristics to film dirt. Concealment of this noise should therefore be possible using the same algorithm. Concealment must take place at the receiver. The addition of a complete motion compensated median filter would be

uneconomic. If the receiver already incorporated DATV bandwidth reduction, however, most of the hardware would already be present. The additional cost of median filtering would then be negligible. Such processing may allow the use of lower signal strengths. This would permit economies in the transmission chain.

Median filtering, while ideal for removing impulse noise, will also reduce broadband noise. Consider a constant signal corrupted by noise. Median filtering the signal will reduce the noise since extreme values are excluded. Let the standard deviation of the signal noise be D the standard deviation of the filtered noise, D', is

D' = 1.2533 D/SQR(N)   Equation 2

where N is the number of filter taps and SQR(N) is the square root thereof. See the book by Kendal and Stuart referenced above, p.251-252. Hence a 3 tap median filter would give a 2.8dB reduction in broadband noise.

The properties of a motion compensated median filter make it suitable for concealing film dirt. A clean, noise free picture is passed substantially unchanged. Noise on an otherwise clean picture is reduced by about 3dB. Positive or negative dirt on monochrome or colour film is replaced by an appropriately shifted (allowing for motion) image from an adjacent picture.

Dirt concealment using this method should have improved performance over currently available techniques.

CLAIMS:

1. A method of processing an input video signal, wherein a plurality of picture-delayed video signals are derived from the input signal, the set of video signals is processed to form a non-linearly temporally filtered output signal corresponding to but containing less impulse noise than the input signal, and the relative delays between the set of video signals are adjusted from area to area of the picture to compensate for movement in the picture.

2. A method according to claim 1, wherein the relative delays are adjusted from pixel to pixel.

3. A method according to claim 1 or 2, wherein the relative delays are adjusted by a process of interpolation compensating for movement within an accuracy which is a fraction of a pixel.

4. A method according to claim 1, 2 or 3, wherein the relative delays are adjusted in dependence upon motion vectors derived from the input video signal.

5. A method according to claim 4, wherein the input signal is a signal derived from a film scanner and the method removes impulse noise resulting from dirt on the scanned film.

6. A method according to claim 1, 2 or 3, wherein the relative delays are adjusted in a DATV receiver in dependence upon motion vectors accompanying the input video signal.

7. A method according to any of claims 1 to 6, wherein the non-linear processing is median filtering.

8. Apparatus for processing an input video signal, comprising a plurality of picture delays for deriving delayed video signals from the input signal, a plurality of interpolators operative upon respective ones of the video signals and responsive motion vectors

describing movement of areas of the picture to generate compensated video signals such that moving objects are brought into register in the compensated video signals and a reference one of the video signals which is not processed by an interpolator, and means operative upon the reference and compensated video signals to produce a non-linearly temporally filtered output signal corresponding to but containing less impulse noise than the input signal.

9. Apparatus according to claim 8, wherein each interpolator comprises a series of cascaded filter taps implementing corresponding segments of an interpolation impulse response.

10. Apparatus according to claim 9, wherein each filter tap comprises a store, means for continually and cyclically writing video data into the store, and means for reading pixels out of the store at addresses offset in accordance with motion vectors pertaining to the pixels and further offset in accordance with offsets individual to the filter taps and corresponding to the relative positions of the said impulse response segments, and the interpolator comprises means for progressively accumulating the read out pixels along the series of filter taps to build up the compensated video signal for that interpolator.

11. Apparatus according to claim 10, wherein each filter tap further comprises a coefficient memory storing the corresponding impulse response segment, means responsive to a fractional part of each motion vector to address the coefficient memory and read out a coefficient, and means for multiplying each pixel read out from the picture store by the currently read out coefficient.

12. Apparatus for processing an input video signal substantially as illustrated in Figs. 3 to 6 of the accompanying drawings.

[illegible] High Holborn, London WC1R 4TP. Further copies may be obtained from The Patent Office,